Low-Rank Matrix Recovery using Gabidulin Codes in Characteristic Zero [1]

Sven Müelich   sven.mueelich@uni-ulm.de
Sven Puchinger   sven.puchinger@uni-ulm.de
Martin Bossert   martin.bossert@uni-ulm.de
Institute of Communications Engineering, University of Ulm, Germany

Abstract. We present a new approach on low-rank matrix recovery (LRMR) based on Gabidulin Codes. Since most applications of LRMR deal with matrices over infinite fields, we use the recently introduced generalization of Gabidulin codes to fields of characteristic zero. We show that LRMR can be reduced to decoding of Gabidulin codes and discuss which field extensions can be used in the code construction.

1   Introduction

Low-rank matrices occur in many applications, e.g., in signal theory, machine learning and collaborative filtering. Unfortunately, in many cases it is only possible to get incomplete or indirect information of matrices. Since applications usually require complete matrices in order to process data, it is necessary to recover matrices from available data. In general this is not possible. However, when matrices are of low-rank, there are efficient algorithms to accomplish this task. So far, *low-rank matrix recovery (LRMR)* is described by a minimization problem which can be solved by convex optimization programs. In this work we show how Gabidulin codes can be used in order to solve the LRMR problem. Section 2 states the LRMR problem, in Section 3 we explain how to generalize Gabidulin codes from finite fields to fields of characteristic zero. In Section 4 we reduce LRMR to decoding of Gabidulin codes. Section 5 gives examples for field extensions which can be used. Finally, Section 6 concludes the paper.

2   Low-Rank Matrix Recovery

Low-Rank Matrix Recovery (LRMR) was first defined in [1, 2, 10] and can be seen as matrix-analogue of compressed sensing [6]. The goal of LRMR is to reconstruct a matrix from incomplete or indirect observations. The problem is stated as follows: We want to recover an unknown matrix $\mathbf{X_0} \in K^{m \times n}$ of lowest possible rank, where in applications K usually is the real or complex field. Therefore, we use observed measurements $\mathbf{y} = \mathcal{A}(\mathbf{X_0})$, which we obtain

---

by applying a linear measurement operator $\mathcal{A} : \mathbb{R}^{m \times n} \to \mathbb{R}^p$ to $\mathbf{X_0}$. We assume that $\mathcal{A}$ can be chosen arbitrarily. Finding a solution to this problem can be specified in terms of the minimization problem

$$\min \mathrm{rank}(\mathbf{X}) \text{ subject to } \mathcal{A}(\mathbf{X}) = \mathcal{A}(\mathbf{X_0}). \tag{1}$$

Often, in literature (1) is also called rank minimization problem. Since this problem is NP-hard, convex relaxations are considered. Most commonly used algorithms are nuclear norm minimization [10], iterative hard thresholding [8], and the alternating projections algorithm [7]. An overview of these and other methods, theoretical guarantees and applications is given in [3].

3 Gabidulin Codes in Characteristic Zero

Gabidulin codes over finite fields were introduced in [4, 5, 13], a comprehensive overview is given in [14]. Since we deal with numbers from infinite alphabets in LRMR, there is a need for Gabidulin codes in characteristic zero, which we introduce according to [9, 11, 12]. Decoding in rank metric can be described by

$$\min \mathrm{rank}(\mathbf{E}') \text{ subject to } \mathbf{HE}' = \mathbf{HE}, \tag{2}$$

which has a similar form as Equation (1). This observation suggests that it might be possible to use a rank metric decoder in order to recover a low-rank matrix and in Section 4 we will show how this can be done.

Let $K \subseteq L$ be fields and L/K be a field extension of degree m. A codeword of a Gabidulin code can either be an $(m \times n)$-matrix over the ground field K or a vector of length n over L. Let $\mathcal{B} = \{\beta_0, \dots, \beta_{m-1}\}$ be a basis of L over K, $\boldsymbol{\beta} = \{\beta_0, \dots, \beta_{m-1}\}$ an order of the basis and $\mathbf{x} \in L$. We use the following bijective mapping in order to switch between the two representations:

Definition 1. *The* ext-*mapping is a bijective function* $\mathrm{ext}_\beta : L^n \to K^{m \times n}$ *related to an ordered basis* $\boldsymbol{\beta}$ *of* L *which fulfills the equation*

$$\mathbf{x} = \beta \cdot \mathrm{ext}_\beta(\mathbf{x}) = \beta \cdot \mathbf{X}. \tag{3}$$

Equation (3) contains the inverse of the mapping. The $\boldsymbol{\beta}$ *in the notation of the mapping* $\mathrm{ext}_\beta(\mathbf{x})$ *is omitted if it is not important to the problem.*

Gabidulin codes in characteristic zero are defined using θ-polynomials which are a generalization of linearized polynomials known from the definition of Gabidulin codes over finite fields.

Definition 2. *Let* $K \subseteq L$ *be fields and* L/K *be a Galois extension. The Galois group of* L/K *is* $\mathrm{Gal}(L/K) = \{\theta : L \to L \text{ automorphism} \mid \theta(k) = k \ \forall k \in K\}$. *The set of* θ-*polynomials is defined as*

$$L[x; \theta] = \left\{ a = \textstyle\sum_{i=0}^{d_a} a_i x^i : a_i \in L, \ d_a \in \mathbb{N}, \ a_{d_a} \neq 0 \right\}.$$

$L[x; \theta]$ with ordinary addition and multiplication $x \cdot \alpha = \theta(\alpha) \; \forall \alpha \in L$ extended to polynomials inductively, is a ring, the so-called θ-polynomial ring. The degree of $a \in L[x; \theta]$ is given by $\deg a = d_a$ and a is called *monic* if $a_{d_a} = 1$. Let $K \subseteq L$ be fields and L/K be a Galois extension of degree m. We denote the number of linearly independent columns over K is by rank_K. We define

$$\mathbf{X}_\theta = \begin{pmatrix} x_1 & \cdots & x_n \\ \vdots & \ddots & \vdots \\ \theta^{s-1}(x_1) & \cdots & \theta^{s-1}(x_n) \end{pmatrix} \text{ and } \mathbf{X}_\mathcal{B} = \begin{pmatrix} x_{1,1} \cdots x_{n,1} \\ \vdots \ddots \vdots \\ x_{1,m} \cdots x_{n,m} \end{pmatrix}.$$

[12, Section 2.2] gives four definitions of rank weight in characteristic zero.

Definition 3. *Let* $\mathbf{x} \in L^n$. *The following are rank weights:*

$$\omega_\mathcal{A}(\mathbf{x}) = deg(\mathcal{A}_{\langle x_1, \ldots, x_n \rangle}), \quad \omega_{\theta,L}(\mathbf{x}) = \mathrm{rank}_L(\mathbf{X}_\theta),$$
$$\omega_{\theta,K}(\mathbf{x}) = \mathrm{rank}_K(\mathbf{X}_\theta), \quad \omega_\mathcal{B}(\mathbf{x}) = \mathrm{rank}_K(\mathbf{X}_\mathcal{B}),$$

where $\mathcal{A}_{\langle x_1, \ldots, x_n \rangle}$ *is the Minimal Subspace Polynomial as defined in [12]. The corresponding rank metrics can be defined as* $\mathrm{d_R}(\mathbf{x}, \mathbf{y}) = \omega_i(\mathbf{x} - \mathbf{y})$.

In the finite field case, these rank weights are the same. In [12, Lemmata 13, 14, and 15], the relation $\omega_\mathcal{A}(\mathbf{x}) = \omega_{\theta,L}(\mathbf{x}) \leq \omega_{\theta,K}(\mathbf{x}) = \omega_\mathcal{B}(\mathbf{x})$ has been proven over characteristic zero.

Definition 4. *Let* $g_1, \ldots, g_n \in L$ *be linearly independent over* K. *Then a Gabidulin code with parameters* n *and* $k \leq n$ *is defined as*

$$\mathcal{C}_\mathrm{G}[n, k] = \{(f(g_1), \ldots, f(g_n)) : f \in L[x; \theta] \wedge \deg f < k\}.$$

For an overview of properties we refer to [12]. For decoding Gabidulin codes in characteristic zero, [11] gives a generalization of a Welch-Berlekamp-like algorithm which allows decoding in $O(n^3)$. In [9] we derived a Gao-like key equation and hence reduced the decoding complexity to $O(n^2)$.

4   New Approach for Low-Rank Matrix Recovery

In this section we reduce the problem of LRMR to decoding of Gabidulin codes and show how decoding can be used in order to recover low-rank matrices. Let $\mathbf{H} \in L^{(n-k)\times n}$ be the parity-check matrix of a Gabidulin code. Recalling Equation (1), we interpret the unknown matrix $\mathbf{X_0} \in K^{m\times n}$ as the low-rank error $\mathbf{E}$ which usually arises while transmitting a codeword of a Gabidulin code. We define the linear measurement operator $\mathcal{A} : K^{m\times n} \to K^p$ as in Algorithm 1.

Algorithm 1: Linear measurement operator

Input: $\mathbf{X_0} \in K^{m \times n}$
Output: $y \in K^p$

1 $\mathbf{x_0} \leftarrow \text{ext}^{-1}(\mathbf{X_0})$;   // $\mathbf{x_0} \in L^n$
2 $\tilde{\mathbf{y}} \leftarrow \mathbf{H} \cdot \mathbf{x_0}^T$;   // $\tilde{\mathbf{y}} \in L^{n-k}$
3 $\mathbf{Y} \leftarrow \text{ext}(\tilde{\mathbf{y}})$;   // $\mathbf{Y} \in K^{(n-k) \times n}$
4 $\mathbf{y} \leftarrow$ vector representation of $\mathbf{Y}$;   // $\mathbf{y} \in K^{n \cdot (n-k)}$
5 **return y**

Note, that all operations used in the algorithm are K-linear.

Theorem 1. *If $p = n \cdot (n - k)$, there exists a mapping $\mathcal{A}$ as in Algorithm 1 such that $\mathbf{X_0}$ with $\text{rank}(\mathbf{X_0}) \leq \frac{n-k}{2} = \frac{p}{2n}$ can be recovered.*

Proof. We use the output of the linear measurement operator to calculate $\tilde{\mathbf{y}} = \text{ext}^{-1}(y) \in L^{n-k}$. This corresponds to the syndrome which is used by a decoder in order to produce the error matrix $\mathbf{E}$ which is in our case the matrix $\mathbf{X_0}$. We know from coding theory, that if $\text{rank}(\mathbf{X_0}) \leq \frac{d-1}{2} = \frac{n-k}{2} = \frac{p}{2n}$ the matrix $\mathbf{X_0}$ (and consequently $\mathbf{E}$) is unique and can be found by a bounded minimum distance decoder [9]. $\square$

Using our result from [9], decoding can be done in $O(n^2)$. In comparison, the complexity of previous methods depends on singular value decomposition (SVD). For instance, nuclear norm minimization [10] is an iterative algorithm, where each iteration requires the computation of a SVD. Which matrices can be recovered by the previous methods depends on a variety of factors like applications and assumptions. There exists rich literature about theoretical guarantees, we refer to the overview provided by [3].

5   Which Fields Can Be Used?

Applications usually use matrices over real or complex numbers. Using Gabidulin codes for solving LRMR, we cannot use $K = \mathbb{C}$, since $\mathbb{C}$ is algebraically closed. It is well-known that the only finite extensions of $\mathbb{R}$ are $\mathbb{C}$ or $\mathbb{R}$ and hence $K = \mathbb{R}$ is also not suitable. Thus, we need another field K which approximates the actual matrix sufficiently well. We can find such a K, e.g., by minimizing the Frobenius norm of the difference matrix: Assuming that $\mathbf{X} \in \mathbb{C}^{m \times n}$ is the actual matrix, we would like to choose a field K such that we can find another matrix $\mathbf{X'} \in K^{m \times n}$ with

$$\|\mathbf{X'} - \mathbf{X}\|_{\mathrm{F}} = \sqrt{\sum_{i,j} |x'_{ij} - x_{ij}|^2} < \epsilon$$

for any $\epsilon > 0$. This condition is always fulfilled if K is a $dense^2$ subfield of $\mathbb{C}$ and helps to minimize the "approximation error" which occurs when going from $\mathbb{C}$ to K.

If we want to recover a matrix over real numbers, we propose to use cyclotomic extensions. Therefore, we take the rational numbers and adjoint the n-th roots of unity. Hence, we get a field extension of degree $\varphi(n)$ [3]. We are able to construct a cyclotomic extension of degree n iff $n \in \text{image}(\varphi)$. Since for every prime p it is $p - 1 \in \text{image}(\varphi)$, the density of $\text{image}(\varphi)$ is at least $\frac{n}{\ln(n)}$ (cmp. prime number theorem). Hence, $\text{image}(\varphi)$ is sufficiently dense to find suitable numbers. The automorphism group of such a field fulfills the required properties, namely the Galois group is cyclic and the characteristic polynomial of an automorphism $\theta \in Gal$ is square-free. Examples can be found in [12, Section 2.3.3].

If the desired matrix is in $\mathbb{C}$, we can use Kummer extensions [12, Section 2.3.3]. We need to choose K to be a dense subfield of $\mathbb{C}$ which has n many n-th roots of unity. We can ensure this by choosing $K = \mathbb{Q}(\zeta_n)$, where $\zeta_n = \mathrm{e}^{\mathrm{i}\frac{2\pi}{n}}$, i.e., K is a cyclotomic extension of $\mathbb{Q}$ of degree $\varphi(n)$. Then, $K \subseteq \mathbb{C}$ dense and it contains n distinct n-th roots of unity (namely ζ_n^i for $i = 0, \ldots, n-1$). Thus, there is a Kummer extension L/K of degree n, e.g., by adjoining an n-th root of 2 to K. This approach works for any $n \in \mathbb{N}$ with $n > 2$.

6   Conclusion

In this work we gave a mathematical description of LRMR and reduced the problem to decoding of Gabidulin codes. Since the matrices in applications usually are over the real or complex field, we used our construction of Gabidulin codes in characteristic zero introduced in [9]. Using this approach, matrices $\mathbf{X}$ with $\text{rank}(\mathbf{X}) \leq \frac{n(n-k)}{2n}$ can be recovered in $O(n^2)$. Finally, we gave ideas of practical fields which can be used in this scenario. In this work we use the case of having exact measurements, however also versions with noise are described in literature, cf. [3, 10]

References

[1] E. Candès, B. Recht, Exact Matrix Completion via Convex Optimization, *Foundations of Computational Mathematics* 9, No. 6, 717-772, 2009.

[2] E. Candès, T. Tao, The Power of Convex Relaxation: Near-optimal Matrix Completion, *IEEE Trans. Inform. Theory* 56, No.5, 2053-2080, 2010.

---

^{2}A field K is dense in $\mathbb{C}$ if for any $z \in \mathbb{C}$ and $\epsilon > 0$, there is an $z' \in K$ with $|z - z'| < \epsilon$.
^{3}Let $\varphi : \mathbb{N} \to \mathbb{N}$ be Euler's phi function known from number theory

[3] M. Davenport, J. Romberg, An Overview of Low-Rank Matrix Recovery from Incomplete Observations, arXiv preprint abs/1601.06422, 2016.

[4] P. Delsarte, Bilinear Forms over a Finite Field, with Applications to Coding Theory, *J. Combin. Theory*, Series A, 25, No. 3, 226-241, 1978.

[5] E. Gabidulin, Theory of Codes with Maximum Rank Distance, *Problemy Peredachi Informatsii* 21, No. 1, 3-16, 1985.

[6] D. Gross, Recovering Low-Rank Matrices from few Coefficients in any Basis, *IEEE Trans. Inform. Theory* 57, No. 3, 1548-1566, 2011.

[7] P. Jain, P. Netrapalli, S. Sanghavi, Low-Rank Matrix Completion using Alternating Minimization, Proc. Forty-fifth annual ACM symposium on Theory of computing, 665-674, 2013.

[8] P. Jain, R. Meka, Raghu, I. Dhillon, Guaranteed Rank Minimization via Singular Value Projection, Advances in Neural Information Processing Systems, 937-945, 2010.

[9] S. Müelich, S. Puchinger, D. Mödinger, M. Bossert, An Alternative Decoding Method for Gabidulin Codes in Characteristic Zero, arXiv preprint arXiv:1601.05205, 2016.

[10] B. Recht, M. Fazel, P. A. Parrilo, Guaranteed Minimum-Rank Solutions of Linear Matrix Equations via Nuclear Norm Minimization, *SIAM review* 52, No. 3, 471-501, 2010.

[11] D. Augot, P. Loidreau, G. Robert, Rank Metric and Gabidulin Codes in Characteristic Zero, ISIT 2013 IEEE International Symposium on Information Theory, 2013.

[12] G. Robert, Codes de Gabidulin en Caractéristique Nulle. Application au Codage Espace-Temps, PhD Thesis, Université Rennes 1, 2015.

[13] R. M. Roth, Maximum-Rank Array Codes and their Application to Crisscross Error Correction, *IEEE Trans. Inform. Theory* 37, No. 2, 328-336, 1991.

[14] A. Wachter-Zeh, Decoding of Block and Convolutional Codes in Rank Metric, PhD Thesis, Université Rennes 1, 2013.